MAC0427

Otimização Não Linear

Segunda Lista de Exercícios

Pedro Gigeck Freire 10737136

① Considere a função $f(x) = x_1^2 + 16x_2^2$ e $\sigma = 0.1$. Calcule um passo do método do gradiente para f com $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e condição de Armijo com σ.

Para calcular a direção de descida, vamos calcular o vetor gradiente ∇f

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 32x_2 \end{pmatrix}$$

Portanto, a direção de descida será

$$d = -\nabla f(x) = \begin{pmatrix} -2x_1 \\ -32x_2 \end{pmatrix}$$

No ponto $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, a direção será $d = \begin{pmatrix} -2 \\ -32 \end{pmatrix}$

Agora, calculamos o passo que satisfaz a condição de Armijo com $\sigma = 0.1$

Ou seja, precisamos encontrar $\lambda > 0$ tal que

$$f(x_0 + \lambda d) \leq f(x_0) + \sigma \lambda \nabla f(x_0)^T d$$

Temos que

- $x_0 + \lambda d = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ -32 \end{pmatrix} = \begin{pmatrix} 1-2\lambda \\ 1-32\lambda \end{pmatrix}$

- $f(x_0 + \lambda d) = f\begin{pmatrix} 1-2\lambda \\ 1-32\lambda \end{pmatrix} = (1-2\lambda)^2 + 16(1-32\lambda)^2 = 1 - 4\lambda + 4\lambda^2 + 16(1 - 64\lambda + 1024\lambda^2)$

$$= 16384\lambda^2 - 1028\lambda + 17$$

• $f(x_0) = 1 + 16 = 17$

• $\sigma \lambda \nabla f(x_0)^T d = 0.1 \lambda \begin{pmatrix} 2 & 32 \end{pmatrix} \begin{pmatrix} -2 \\ -32 \end{pmatrix} = \lambda\, 0.1 \,(-4 - 1024) = \lambda \dfrac{(-1028)}{10}$

Portanto,

$$f(x_0 + \lambda d) \leq f(x_0) + \sigma \lambda \nabla f(x_0)^T d \iff$$

$$16384\lambda^2 - 1028\lambda + 17 \leq 17 + \lambda \frac{(-1028)}{10} \iff$$

$$16384\lambda^2 \leq \lambda\left(\frac{-1028}{10} + 1028\right) \iff$$

$$16384\lambda \leq \frac{9}{10} 1028 \qquad (\text{pois } \lambda > 0) \iff$$

$$\lambda \leq \frac{9}{10} \frac{1028}{16384} \cong 0,05646\ldots$$

Logo, qualquer $\lambda \in (0, 0,0564]$ satisfaz a condisão de Armijo. Para simplificar, podemos tomar $\lambda = 0,05 = \frac{1}{20}$

Assim, o primeiro passo do método do gradiente obterá

$$x_1 = x_0 + \lambda d = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.05 \begin{pmatrix} -2 \\ -32 \end{pmatrix} = \begin{pmatrix} 1 - 0.1 \\ 1 - 0.16 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.84 \end{pmatrix}$$

Com $f(x_1) = (0.9)^2 + 16(0.84)^2 = 0,81 + 16 \cdot 0,7056 = 0,81 + 11,2896 = 12,0996$

(2) Considere a funsão $f(x) = x_1^2 + 2x_1x_2 + 4x_2^2 + x_2x_3 + 16x_3^2$ e $\sigma = 0.1$. Calcule um passo do método do gradiente para f com $x_0 = (1, 1, 1)^T$ e condisão de Armijo com σ.

Repetindo os passos do exercício anterior, vamos calcular o gradiente de f

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \frac{\partial f}{\partial x_3}(x) \end{pmatrix} = \begin{pmatrix} 2x_1 + 2x_2 \\ 2x_1 + 8x_2 + x_3 \\ x_2 + 32x_3 \end{pmatrix}$$

Assim,

$$\nabla f(x_0) = \nabla f\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right) = \begin{pmatrix}2+2\\2+8+1\\1+32\end{pmatrix} = \begin{pmatrix}4\\11\\33\end{pmatrix}$$

E a direção de descida será

$$d = \begin{pmatrix}-4\\-11\\-33\end{pmatrix}$$

Para encontrar λ que satisfaça a condição de Armijo, calculamos os termos separadamente:

- $x_0 + \lambda d = (1,1,1)^T + \lambda(-4,-11,-33)^T = (1-4\lambda,\ 1-11\lambda,\ 1-33\lambda)^T$

- $f(x_0 + \lambda d) = f((1-4\lambda,\ 1-11\lambda,\ 1-33\lambda)^T)$

$$= (1-4\lambda)^2 + 2(1-4\lambda)(1-11\lambda) + 4(1-11\lambda)^2 + (1-11\lambda)(1-33\lambda) + 16(1-33\lambda)^2$$

$$= 1 - 8\lambda + 16\lambda^2 + 2(1 - 11\lambda - 4\lambda + 44\lambda^2) + 4(1 - 22\lambda + 121\lambda^2) + (1 - 44\lambda + 363\lambda^2) + 16(1 - 66\lambda + 1089\lambda^2)$$

$$= 1 - 8\lambda + 16\lambda^2 + 2 - 30\lambda + 88\lambda^2 + 4 - 88\lambda + 484\lambda^2 + 1 - 44\lambda + 363\lambda^2 + 16 - 1056\lambda + 17424\lambda^2$$

$$= 18375\lambda^2 - 1226\lambda + 24$$

- $f(x_0) = f((1,1,1)^T) = 1 + 2 + 4 + 1 + 16 = 24$

- $\sigma\lambda\nabla f(x_0)^T d = 0.1\lambda\,(4,11,33)(-4,-11,-33)^T = \lambda\, 0.1\,(-16 - 121 - 1089) = \lambda\dfrac{(-1226)}{10}$

Portanto, para λ satisfazer a condição de Armijo com $\sigma = 0.1$, temos

$$f(x_0 + \lambda d) \leq f(x_0) + \sigma\lambda\nabla f(x_0)^T d \iff$$

$$18375\lambda^2 - 1226\lambda + 24 \leq 24 + \lambda\frac{(-1226)}{10} \iff$$

$$18375\lambda^2 \leq \lambda\frac{9}{10}1226 \iff \quad \text{pois } \lambda > 0$$

$$\lambda \leq \frac{9}{10}\,\frac{1226}{18375} \cong 0{,}06004$$

Logo, todo $\lambda \in [0, 0{,}06]$ satisfaz a condição de Armijo

Podemos tomar, portanto, o primeiro passo do método do gradiente $\boxed{\lambda = 0.06}$

Obtendo

$$\underline{x_1} = x_0 + \lambda d = (1,1,1)^T + 0.06\,(-4,-11,-33)^T = (1-0{,}24,\ 1-0{,}66,\ 1-1{,}98)^T$$

$$= \underline{(0.76,\ 0.44,\ -0.98)}$$

Com

$$\underline{f(x_1)} = (0.76)^2 + 2(0.76 \cdot 0.44) + 4(0.44)^2 + (0.44)(-0.98) + 16(-0.98)^2 =$$

$$= 0{,}5776 + 0{,}6688 + 0{,}7744 + (-0{,}4312) + 15{,}3664 = \underline{16{,}956}$$

3 Sejam $f: \mathbb{R}^n \to \mathbb{R}$ com derivadas parciais contínuas, $x, d \in \mathbb{R}^n$ e $\lambda > 0$ tais que $x + \lambda d$ satisfaz a condição de Armijo com constante σ. Seja $\mu \in (0, \lambda)$. $x + \mu d$ satisfaz a mesma condição de Armijo? Prove ou dê um contra exemplo.

Não. $x + \mu d$ não necessariamente satisfaz a condição de Armijo.

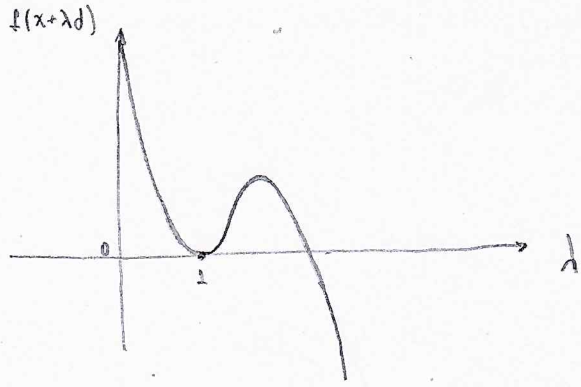
Contra exemplo:

Considere $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x_1, x_2) = x_1^2 - x_1^3$$

Tome $x = (-1, 0)$ e $d = (1, 0)$. Considere $\sigma = 0{,}25 = 1/4$.

Esboçando a função (de uma variável) $f(x + \lambda d) = f((\lambda - 1, 0)) = (\lambda-1)^2 - (\lambda-1)^3$



Agora, vamos colocar no gráfico o lado esquerdo da desigualdade da condisão de Armijo

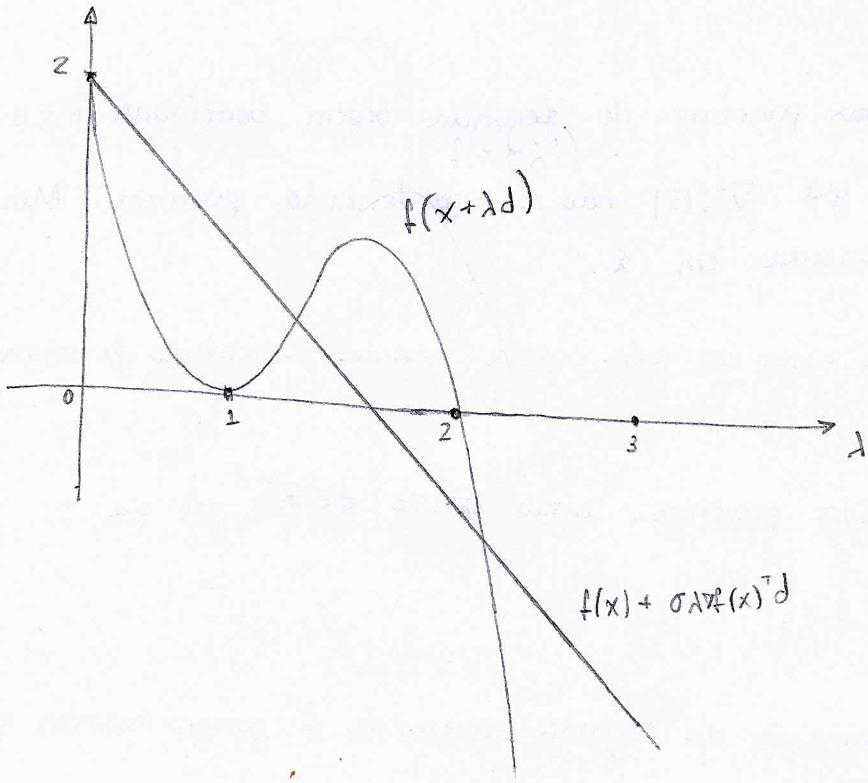
$$f(x) + \sigma\lambda \nabla f(x)^T d$$

Antes, calculemos o gradiente no ponto $x = (-1, 0)$

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \end{pmatrix} = \begin{pmatrix} 2x_1 - 3x_1^2 (x) \\ 0 \end{pmatrix} = \begin{pmatrix} -2-3 \\ 0 \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \end{pmatrix}$$

Portanto $\quad f(x) + \sigma\lambda \nabla f(x)^T d = (-1)^2 - (-1)^3 + \frac{1}{4}\lambda\,(-5\;\;0)(1\;\;0)^T$

$$= 2 + \lambda(-5) = 2 - \frac{5}{4}\lambda$$



Portanto, se tomarmos $\lambda = 4$, a condisão de Armijo será satisfeita:

$$f(x+4d) \leqslant f(x) + \sigma\, 4\, \nabla f(x)^T d \iff$$

$$f((-1+4, 0)) \leqslant f((-1,0)) + 4/4\,(-5,0)(1,0)^T \iff$$

$$(-5)^2 - (-5)^3 \leqslant (-1)^2 - (-1)^3 + 1.(-5) \iff$$

$$25 - 125 \leqslant 1 - 1 - 5 \iff$$

$$-100 \leqslant -5 \qquad \checkmark \qquad \text{(de fato, satisfaz)}$$

Porém, quando tomamos $\mu = 2$; a condisão de Armijo não é satisfeita:

$$f(x+2d) \leqslant f(x) + \sigma 2 \nabla f(x)^T d \iff$$

$$f((-1+2, 0)) \leqslant (-1)^2 - (-1)^3 + \frac{2}{4}(-5 \;\; 0)(1 \;\; 0)^T \iff$$

$$1^2 - 1^3 \leqslant 1 - 1 + \frac{1}{2}(-5) \iff$$

$$0 \leqslant -5/2 \qquad \underline{\text{FALSO}}$$

De fato, a condisão de Armijo é satisfeita com $\lambda = 4$ mas não é satisfeita com $\mu = 2$.

4 Sejam $f: \mathbb{R}^n \to \mathbb{R}$ com derivadas parciais de segunda ordem contínuas e $\bar{x} \in \mathbb{R}^n$ tal que $\nabla f(\bar{x}) = 0$ e suponha que $\nabla^2 f(\bar{x})$ não é semidefinida positiva. Mostre que existe uma diresão de descida em $\bar{x}$.

Vamos utilizar a mesma ideia vista em aula para provar a condisão de otimalidade de segunda ordem:

Como $\nabla^2 f(\bar{x})$ não é semidefinida positiva, então existe $d \in \mathbb{R}^n$ tal que

$$d^T \nabla^2 f(\bar{x}) d < 0 .$$

Agora podemos utilizar a aproximasão de segunda ordem de f para mostrar que d é uma diresão de descida

Temos que

$$f(\bar{x}+d) = f(\bar{x}) + \nabla f(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f(\bar{x}) d + o(\|d\|^2)$$

$$= f(\bar{x}) + \frac{1}{2} d^T \nabla f(\bar{x}) d + o(\|d\|^2) \quad (\text{pois } \nabla f(\bar{x}) = 0)$$

$$< f(\bar{x}) + o(\|d\|^2) \quad (\text{pois } d^T \nabla f(\bar{x}) d < 0)$$

Portanto, para $\|d\|$ suficientemente pequeno, teremos que

$f(\bar{x} + d) < f(\bar{x})$

Logo, d é uma direção de descida.

⑤ Sejam $f: \mathbb{R}^n \to \mathbb{R}$ e $\bar{x} \in \mathbb{R}^n$ tal que $\nabla f(\bar{x}) \neq 0$. Se $M \in \mathbb{R}^{n\times n}$ é positiva definida mostre que $d = -M\nabla f(\bar{x})$ é uma direção de descida.

Vamos usar a aproximação de 1ª ordem para $f(\bar{x} + (-M\nabla f(\bar{x})))$

$$f(\bar{x} - M\nabla f(\bar{x})) = f(\bar{x}) + \nabla f(\bar{x})^T(-M\nabla f(\bar{x})) + o(\|-M\nabla f(\bar{x})\|)$$

$$= f(\bar{x}) - \nabla f(\bar{x})^T M \nabla f(x) + o(\|-M\nabla f(\bar{x})\|)$$

$$< f(\bar{x}) + o(\|-M\nabla f(x)\|)$$ pois $\nabla f(\bar{x})^T M \nabla f(\bar{x}) > 0$ já que M é positiva definida.

Logo, multiplicando $-M\nabla f(\bar{x})$ por uma constante λ suficientemente pequena, teremos

$$f(\bar{x} - \lambda M\nabla f(\bar{x})) < f(\bar{x})$$

Portanto $-M\nabla f(\bar{x})$ é uma direção de descida.

⑥ Considere o método do gradiente com busca linear exata, isto é, x_{k+1} é o minimizador local de f ao longo da reta $\{x_k - t\nabla f(x_k),\ t \in \mathbb{R}\}$. Prove que $\nabla f(x_{k+1})^T \nabla f(x_k) = 0$.

Considere f ao longo da reta $\{x_k - t\nabla f(x_k) : t \in \mathbb{R}\}$, isto é, seja

$\varphi: \mathbb{R} \to \mathbb{R}$ dada por $\varphi(t) = f(x_k - t\nabla f(x_k))$

Note que

$$\varphi'(t) = \nabla f(x_k - t\nabla f(x_k))^T(-\nabla f(x_k))$$

Agora consideremos um ponto t^* de mínimo local de f.

Como t^* é mínimo local, então $\varphi'(t^*) = 0$

Além disso, como x_{k+1} é o ponto de mínimo local de f ao longo da reta $\{x_k - t\nabla f(x_k) : t \in \mathbb{R}\}$, então

$$x_{k+1} = x_k - t^* \nabla f(x_k)$$

Portanto

$$0 = \varphi'(t^*) = \nabla f(x_k - t^* \nabla f(x_k))^T (-\nabla f(x_k))$$

$$= -\nabla f(x_{k+1})^T \nabla f(x_k)$$

De fato, $\nabla f(x_{k+1})^T \nabla f(x_k) = 0$

7) Seja $f: \mathbb{R}^n \to \mathbb{R}$ diferenciável em $\bar{x}$ e sejam $d^1, \dots, d^n \in \mathbb{R}^n$ vetores linearmente independentes. Suponha que o mínimo de $g_j(\lambda) = f(\bar{x} + \lambda d^j)$ ocorra em $\lambda = 0$ para $j = 1, \dots, n$. Prove que $\nabla f(\bar{x}) = 0$. isso implica que f tem um mínimo local em $\bar{x}$?

Note que, para todo $j = 1, 2, \dots, n$

$$g_j'(\lambda) = \nabla f(\bar{x} + \lambda d^j)^T d^j$$

E que $g_j'(0) = 0$ pois $\lambda = 0$ é mínimo de g_j.

Então

$$0 = g_j'(0) = \nabla f(\bar{x})^T d^j.$$

Podemos somar todas as $g_j'(0)$ e obter

$$0 = \sum_{j=1}^{n} g_j'(0) = \sum_{j=1}^{n} \nabla f(\bar{x})^T d^j = \nabla f(\bar{x})^T \left(\sum_{j=1}^{n} d^j \right)$$

Isso implica que $\nabla f(\bar{x}) = 0$ pois d^j são linearmente independentes, portanto os coeficientes da soma têm que ser todos zero.

Vamos verificar que f tem mínimo local em $\bar{x}$:

Note que, para todo $j = 1, 2, \dots, n$

$$g_j''(\lambda) = \left[\nabla^2 f(\bar{x} + \lambda d^j)^T d^j\right]^T d^j$$

$$= d^{jT} \nabla^2 f(\bar{x} + \lambda d^j)\, d^j$$

E como $\lambda = 0$ é mínimo local de g_j, temos que

$$d^{jT} \nabla^2 f(\bar{x})\, d^j = g_j''(0) \geq 0$$

Logo, como $\{d^j : j = 1, \dots, n\}$ é um conjunto linearmente independente, temos que é uma base de $\mathbb{R}^n$. Então todo vetor $x \in \mathbb{R}$ pode ser decomposto em uma combinação linear dos d^j, $x = a_1 d_1 + \dots + a_n d_n$ $(a_j \in \mathbb{R}^n, j = 1, \dots, n)$.

Portanto $x^T \nabla^2 f(\bar{x})\, x = (a_1 d^1 + \dots + a_n d^n) \nabla^2 f(\bar{x}) (a_1 d^1 + \dots + a_n d^n)$

$$= (a_1 d^1)^T \nabla^2 f(\bar{x}) (a_1 d^1) + \dots + (a_n d^n)^T \nabla^2 f(\bar{x}) (a_n d^n)$$

$$= a_1^2\, d^{1T} \nabla^2 f(\bar{x})\, d^1 + \dots + a_n^2\, d^{nT} \nabla^2 f(\bar{x})\, d^n$$

$$= a_1^2\, g_1''(0) + \dots + a_n^2\, g_n''(0)$$

≥ 0 pois $a_j^2 \geq 0$ e $g_j''(0) \geq 0$

Isto é, $\nabla^2 f(\bar{x})$ é uma matriz semidefinida positiva. Isso implica que f tem mínimo local em $\bar{x}$.

8) Seja

$$f(x) = \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{2}(1 - x_1)^2$$

Qual é o minimizador de f? Faça uma iteração do método de Newton para f a partir de $x_0 = (2, 2)^T$ com $\lambda = 1$. É um bom passo? Antes de decidir, calcule $f(x_0)$ e $f(x_1)$.

- Vamos calcular o mínimo de f:

Temos que

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \\ \frac{\partial f}{\partial x_2}(x) \end{pmatrix} = \begin{pmatrix} (x_1 - x_2) - (1 - x_1) \\ \\ -(x_1 - x_2) \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 - 1 \\ \\ x_2 - x_1 \end{pmatrix}$$

De modo que

$$\nabla f(x) = 0 \iff \begin{cases} 2x_1 - x_2 - 1 = 0 \\ x_2 - x_1 = 0 \end{cases} \iff x_1 = x_2 = 1$$

Como $f(1,1) = 0$ e $f(x) \geq 0$ para todo $x \in \mathbb{R}^2$, então

$\underline{\bar{x} = (1,1)^T}$ é minimizador de f.

- Agora, vamos fazer uma iteração do método de Newton, com $x_0 = (2,2)^T$ e $\lambda = 1$.

Antes, calculemos $\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ para todo $x \in \mathbb{R}^n$

E a inversa é $\nabla^2 f(x)^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

Então a direção de descida será

$$-\nabla^2 f(x_0)^{-1} \nabla f(x_0) = -\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 4-2-1 \\ 2-2 \end{pmatrix} = -\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Com o passo $\lambda = 1$, obteremos

$$x_1 = x_0 + \lambda d = (2,2)^T + (-1,-1)^T = (1,1)^T$$

Que é o mínimo da função!

Portanto o método precisou de apenas uma iterasão para chegar na resposta.

Na primeira iterasao tínhamos $f(x_0) = \frac{1}{2}$

Na segunda, teremos $f(x_1) = 0$.

Ou seja, $\lambda = 1$ foi um bom passo, nesse caso.

9 Obtenha uma base para o núcleo da Matriz

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$$

e encontre $x \in \mathbb{R}^4$ tal que $Ax = (1, 2)^T$

Vamos calcular o escalonamento de A

$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$ subtraimos 5 vezes a linha 1 da linha 2

$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \end{pmatrix}$ Agora multiplicamos a linha 2 por $-\frac{1}{4}$

$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \end{pmatrix}$ Por fim, subtraimos 2 vezes a linha 2 da linha 1

$\begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix}$

Agora, podemos calcular o núcleo da matriz escalonada
Teremos

$$\begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0 \iff \begin{cases} x - z - 2w = 0 \\ y + 2z + 3w = 0 \end{cases} \iff \begin{cases} x = z + 2w \\ y = -2z - 3w \end{cases}$$

Então $x = (z+2w, -2z-3w, z, w)^T = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix} w$

Ou seja, os vetores $\begin{pmatrix}1\\-2\\1\\0\end{pmatrix}$ e $\begin{pmatrix}2\\-3\\0\\1\end{pmatrix}$ formam uma base do núcleo de A.

- Por fim, vamos resolver o sistema $Ax = \begin{pmatrix}1\\2\end{pmatrix}$

$$\begin{pmatrix}1 & 2 & 3 & 4\\5 & 6 & 7 & 8\end{pmatrix}\begin{pmatrix}x\\y\\z\\w\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix} \iff \begin{cases} x + 2y + 3z + 4w = 1 \\ 5x + 6y + 7z + 8w = 2 \end{cases}$$

E podemos tomar $z = w = 0$ e obter o sistema $\begin{cases} x + 2y = 1 \\ 5x + 6y = 2 \end{cases}$

$$\iff \begin{cases} 5x + 10y = 5 \\ 5x + 6y = 2 \end{cases} \iff 4y = 3 \iff y = 4/3 \quad , \quad x = 2y - 1 = 8/3 - 1 = 5/3$$

Obtemos, portanto, a solução $\begin{pmatrix}5/3\\4/3\\0\\0\end{pmatrix}$

(10) Obtenha uma base para o núcleo da matriz

$$A = \begin{pmatrix}1 & 2 & 3 & 4\\5 & 6 & 7 & 8\\7 & 10 & 13 & 16\end{pmatrix}$$

e mostre que não existe $x \in \mathbb{R}^4$ tal que $Ax = (1, 2, 3)^T$

Primeiramente, escalonamos a matriz A: (partindo já da versão escalonada do exercício anterior:

$$A = \begin{pmatrix}1 & 0 & -1 & -2\\0 & 1 & 2 & 3\\7 & 10 & 13 & 16\end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 10 & 20 & 30 \end{pmatrix}$$ subtraindo 7 vezes a primeira linha da terceira

$$\begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$ subtraindo 10 vezes a segunda linha da última

Portanto, vimos que a matriz A não tem posto máximo

Achando uma base para o núcleo, obteremos a mesma do exercício 9:

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Agora, podemos provar que não existe x tal que $Ax = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ mostrando que esse vetor $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ não é perpendicular ao núcleo, ou seja, não está na imagem de A.

Obtemos, da forma escalonada, que as colunas 1 e 2 da matriz A formam uma base do espaso coluna.

Portanto o espaso coluna é gerado por

$$\begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix} x + \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix} y \quad (x, y \in \mathbb{R})$$

E conferimos que o sistema

$$\begin{cases} x + 2y = 1 \\ 5x + 6y = 2 \\ 7x + 10y = 3 \end{cases}$$ Não tem solusão.